Comparative effectiveness research priorities: Identifying critical gaps in evidence for clinical and health policy decision making

Kalipso Chalkidou
National Institute of Health and Clinical Excellence

Danielle Whicher
Center for Medical Technology Policy

Weslie Kary
Integrated Healthcare Association

Sean Tunis
Center for Medical Technology Policy

Background: In the debate on improving the quality and efficiency of the United States healthcare system, comparative effectiveness research is increasingly seen as a tool for reducing costs without compromising outcomes. Furthermore, the recent American Recovery and Reinvestment Act explicitly describes a prioritization function for establishing a comparative effectiveness research agenda. However, how such a function, in terms of methods and process, would go about identifying the most important priorities warranting further research has received little attention.
Objectives: This study describes an Agency for Healthcare Research and Quality-funded pilot project to translate one current comparative effectiveness review into a prioritized list of evidence gaps and research questions reflecting the views of the healthcare decision makers involved in the pilot.
Methods: To create a prioritized research agenda, we developed an interactive nominal group process that relied on a multistakeholder workgroup scoring a list of research questions on the management of coronary artery disease.
Results: According to the group, the areas of greatest uncertainty regarding the management of coronary artery disease are the comparative effectiveness of medical therapy versus percutaneous coronary interventions versus coronary artery bypass grafting for different patient subgroups; the impact of diagnostic testing; and the most effective method of developing performance measures for providers.

Conclusions: By applying our nominal group process, we were able to create a list of research priorities for healthcare decision makers. Future research should focus on refining this process because determining research priorities is essential to the success of developing an infrastructure for comparative effectiveness research.

Keywords: Investigational therapies, Healthcare reform, Health services research, Coronary artery disease

COMPARATIVE EFFECTIVENESS RESEARCH AS A POLICY OPTION FOR IMPROVING QUALITY AND REDUCING WASTE

In the debate on improving the quality and efficiency of the United States healthcare system, comparative effectiveness research (CER) is increasingly cited as a possible tool for reducing costs without compromising outcomes. Numerous recent policy reports, academic publications, and legislative bills have tried to define, operationalize, and estimate the potential impact of using CER to inform health policy and practice in the United States (2;8;12–15;19).

At the same time, several bills have been considered in both the House of Representatives and the Senate, with the Comparative Effectiveness Research Act of 2008 being the most recent one, calling for the launch of a dedicated institute to undertake CER in the United States (20).

In addition, Congress recently approved the American Recovery and Reinvestment Act, which included $1.1 billion for CER to be shared between the National Institutes of Health (NIH), the Agency for Healthcare Research and Quality (AHRQ), and the department of Health and Human Services (HHS). Approximately $1 million is to be allocated to the Institute of Medicine (IOM) "to produce and submit a report to the Congress and the Secretary by no later than June 30, 2009, that includes recommendations on the national priorities for comparative effectiveness research to be conducted or supported with the [HHS] funds . . . and that considers input from stakeholders." At the same time the multistakeholder Federal Coordinating Council, set up to provide coordination and strategic input on organizational needs for CER, is also charged with supporting the prioritization process (3).

However, while there is clearly an increasing level of support for CER from public and private sector stakeholders, the question of how, in terms of both methods and process, to go about identifying the most important priorities warranting further research has received little attention, apart from a few exceptions. In 2007, the Medicare Evidence Development and Coverage Advisory Committee (MedCAC) convened a group of scientists to create a list of evidentiary gaps important to the Medicare population (14). In addition, the IOM's Priority Assessment Inventory Project aims to identify, using expert opinion, a list of priorities for CER, which would then inform policy and treatment decisions (11). Furthermore, AHRQ has done a substantial amount of work on priority

setting methods for topics for comparative effectiveness reviews in addition to supporting comparative effectiveness reviews into key areas of uncertainty (1).

However, there is still a lot of work that needs to be done to develop a method for identifying and prioritizing those research areas where the $1.1 billion should be invested.

IDENTIFYING HIGH PRIORITIES FOR COMPARATIVE EFFECTIVENESS RESEARCH

The public and private investment in clinical research in the United States is significant. Each year, the NIH spends $30 billion on trials, whereas another $58.8 billion was spent on commercial research and development in 2007 (16). At the same time, over 12,000 new clinical trials are published every year (18). However, most systematic reviews and practice guidelines continue to point out the lack of evidence and call for more research into key areas for informing decision making in health care. To the extent CER is meant to address specific questions around everyday policy and practice, it may be that the noted inadequacy of the available evidence reflects a failure of current private and public research arrangements to identify, articulate, and address the questions that really matter to decision makers.

In 2007, AHRQ funded the Center for Medical Technology Policy (CMTP) to conduct a pilot project entitled Priorities for Evidence Development (PED) to address some of the aforementioned issues. The project had two deliverables: first, a prioritized list of research questions reflecting gaps in the evidence in a selected clinical disease area; second, the development of a reproducible, methodologically robust, user-friendly framework for identifying and prioritizing critical knowledge gaps, applicable to different clinical areas and types of evidence synthesis. AHRQ was the main customer of this work and their main objective was to develop questions that might prove appropriate for CER. An additional potential end user of the high priority research list are other funding agencies and stakeholders interested in supporting research into key evidence gaps in the field.

Below, we give an overview of the methods and process followed during the pilot, describe our findings, and discuss the strengths and weaknesses of our approach while drawing lessons regarding how the process could be improved in the context of the current drive to boost CER capacity in the United States.

METHODS

Priorities for Evidence Development: Project Overview

The objective of this project was to translate one current comparative effectiveness review into a prioritized list of research questions reflecting the views of the healthcare decision makers. Multistakeholder engagement was the core attribute of the overall process. A detailed description of the process to generate these research priorities is given below.

Selecting a Comparative Effectiveness Report.

The 2007 EPC report *Comparative Effectiveness Review of Percutaneous Coronary Interventions (PCI) and Coronary Artery Bypass Graft Surgery (CABG)* (5) was selected as the starting point for this project. Several criteria informed our selection of this specific review, including the following: (i) Timing; this was a recently concluded report and thus considered to be a topical and up-to-date review; (ii) Amenability of topic to CER; (iii) Inclusion of one or more new technologies that have potential for fast diffusion and high cost impact across the healthcare system; (iv) Inclusion of one or more new technologies that have potential for bringing about important health gains to the population; (v) Existence of previous or ongoing initiatives harvesting patient and consumer input in this field; and (vi) Consistency with overall AHRQ priorities.

Conflicting, often inconclusive, evidence on the different treatment options for coronary artery disease (CAD) coupled with high diffusion rates of emerging PCI technologies make this an ideal topic for undertaking further CER to inform more effective clinical practice. Furthermore, within this clinical area, there are currently wide geographic variations in practice patterns across the United States, and many unresolved questions related to everyday practice remain, with the potential to impact healthcare outcomes, particularly among populations served by the Centers for Medicare and Medicaid Services (CMS).

Collating the Evidence Gaps and Convening a Multistakeholder Working Party.

After selecting a comparative effectiveness review, a multistakeholder workgroup, comprising representatives from hospitals, payers, product manufacturers, clinicians, researchers, consumers, and government agencies, was convened to undertake the prioritization process (Table 1). Seats were allocated among the relevant stakeholders so that there was some degree of proportionality between different interest groups—a factor that was particularly important for the voting stage of the prioritization process, where no weights were applied to individual votes.

The workgroup was convened on three occasions throughout the project to help identify, review, refine, and prioritize the evidence gaps and, in parallel, to help develop, in an iterative manner, the actual prioritization criteria and

Table 1. Composition of Working Party

Stakeholder group	No. of participants
Hospitals	2
Payers	3
Product industry	3
Clinicians	5
Researchers	3
Patients/consumers	3
Government agencies	2
Total	21

Note. The multistakeholder Working Party convened by Center for Medical Technology Policy was asked to score the importance each of the priority research questions identified in the preliminary stages of this project. No weights were applied to individual votes.

process. Additional input was provided by workgroup members between meetings, as required.

Initially, CMTP staff, workgroup members, and project consultants with expertise in this clinical area generated a list of gaps in current evidence surrounding the clinical use of PCI and CABG using the discussion section of the PCI versus CABG EPC report (5) as a starting point. Information from other current systematic reviews and evidence-based policy documents, peer-reviewed publications, and reviews of published studies were also incorporated into the list of evidentiary gaps, as were recommendations from expert consultants, including the authors of the review and workgroup members.

Transforming Identified Research Gaps into Researchable Questions.

Once identified, the evidence gaps were translated into researchable questions, again through several iterations between group members, with support from our expert consultants who had experience in trial design. A list of the evidence gaps is given as an online supplement (Supplementary Item 1, which can be viewed online at www.journals.cambridge.org/thc).

The Prioritization Process.

There are several different approaches to prioritization, including formal consensus methods such as Delphi and nominal group techniques and economic impact approaches such as the payback approach or expected value of information models (6;17;21). Having considered the various alternatives, we opted for an informal nominal group prioritization process based on an explicit set of criteria pre-agreed upon by the group. Similar approaches have been widely tested and adopted in the United States and abroad (4;7;9;10) and tend to have less burdensome resource implications (such as the need to use "expert judges/facilitators" or to undertake additional economic analyses).

A set of priority setting criteria were developed in collaboration with the workgroup and served as general guiding factors when prioritizing: (i) Impact on patient health/outcomes, and the intervention effectiveness compared to available alternatives; (ii) Current and projected use of the intervention:

variation in practice and diffusion rates; (iii) Safety concerns; (iv) Quantity and quality of the research so far including systematic reviews and research currently planned or in progress; (v) Most appropriate research design and feasibility of research, including costs, randomization issues, and timing, particularly in relation to fast evolving or diffusing technologies; and (vi) Uncertainty surrounding the use of the intervention, particularly in population subgroups (e.g., by age, gender, ethnicity, comorbidities, and so on).

Workgroup members were asked to score each research question using a scale of 3 to 1 (with 3 representing higher and 1 representing lower importance) by integrating the suggested criteria in a qualitative manner, where appropriate, rather than scoring each question against each criterion. The prioritization process consisted of two rounds of scoring. After seeing the results of the first round of scoring, participants discussed their scores and the resulting rankings. This discussion was followed by a second round of scoring, during which workgroup members were provided with additional information on relevant ongoing or recently reported research related to the evidence gaps under consideration (Supplementary Item 2, which can be viewed online at www.journals.cambridge.org/thc) (23). All voting took place anonymously.

After each round of scoring, the mean and median score for each question were calculated and the research questions were ranked, with those receiving the highest mean priority score at the top of the list. Additionally, the mean deviation from the median score was calculated for each of the questions to assess the degree of agreement among the workgroup members. The mean and median were also calculated separately for each of the stakeholder groups to determine if and how stakeholder groups differed in their responses. A Wilcoxon rank sum test was performed to determine if the degree of change in the overall rankings between the first and second rounds of scoring was significant.

Consultation with Professional Organizations.

After the end of the process, the input of professional societies was sought on the final list of prioritized questions, including the American College of Cardiologists (ACC), the Society of Thoracic Surgeons (STS), and the Society for Cardiovascular Angiography and Interventions (SCAI). We draw on these additional comments in the discussion section of the study.

RESULTS

Top Priorities

Scoring sheets containing the list of twenty-three researchable questions were circulated for the first round of voting (Supplementary Item 3, which can be viewed online at www.journals.cambridge.org/thc); the response rate for this first round was 90.5% (19 of 21 workgroup members). Fourteen questions received mean scores of 2 or above and were

Table 2. Top Five Research Priorities following Second Round of Scoring

Rank	Question
1	What is the comparative effectiveness of medical therapy vs. PCI vs. CABG in terms of reducing adverse objective outcomes and improving subjective outcomes in patients with CAD over 75 years of age, women, and ethnic minorities?
2	What is the impact of diagnostic tests (invasive, noninvasive catheterization) used on patients with suspected CAD on treatment modality and final outcomes? Consider invasive and noninvasive cardiac catheterization such as exercise treadmill test (ETT), echo ETT, CT angiography and coronary calcification, MRI angiography. Can emerging noninvasive diagnostic tests replace invasive catheterization?
3	How can we develop evidence-based sufficiently-sensitive performance measures for PCI and CABG (other than volume) to identify/predict institutions or physicians with poor quality CABG and PCI outcomes?
4	What is the comparative effectiveness of medical therapy vs. PCI vs. CABG in terms of reducing adverse objective outcomes and improving subjective outcomes in patients with different extent of CAD based on angiography (two- vs. three-vessel)?
5	What is the comparative effectiveness of medical therapy vs. PCI vs. CABG in terms of reducing adverse objective outcomes and improving subjective outcomes in patients with CAD and heart failure?

Note. The questions above reflect the top five research priorities concerning the comparative clinical effectiveness of Percutaneous Coronary Interventions versus Coronary Artery Bypass Grafting for the treatment of coronary artery disease based on an iterative scoring process by Working Party members.
PCI, percutaneous coronary intervention; CABG, coronary artery bypass graft; CAD, coronary artery disease; CT, computed tomography; MRI, magnetic resonance imaging.

selected to be included in the second round of scoring by the workgroup. The workgroup discussed the top fourteen research questions during its second teleconference. After this meeting, workgroup participants were provided with additional information on relevant research, including any relevant planned, on-going, or recently completed studies addressing the top questions (Supplementary Item 3, which can be viewed online at www.journals.cambridge.org/thc) and were asked to re-score those questions. The resulting top five priorities (Table 2) after the second round of scoring were selected for further discussion at the final workgroup meeting.

The level of agreement between the group members tended to increase for the top five research priorities as the deliberations and voting cycles progressed. The major changes in the ordering of individual questions between the two rounds of voting seemed to be due to discussions after the first round of voting of relevant ongoing trials that could address some of the questions considered to be of high priority (e.g., ongoing research into patients with CAD and

diabetes), feedback to the group by specialist and researcher members of the workgroup on the rates of diffusion, variation in practice, and increased uncertainty around specific interventions (e.g., noninvasive imaging tests), and the severe limitations of currently available performance measures such as mortality or volume of procedures.

According to the group's scoring results, the areas of greatest uncertainty regarding the management of CAD were the comparative effectiveness of medical therapy versus PCI versus CABG for different patient subgroups, the impact of diagnostic testing, and the most effective method of developing performance measures for providers. There was general agreement among the stakeholder groups as to the highest priority questions in the second round.

Input from Professional Society Representatives

In addition to engaging the workgroup, we sought feedback on the project methodology and results from representatives from the American College of Cardiologists (ACC), the Society of Thoracic Surgeons (STS), and the Society for Cardiovascular Angiography and Interventions (SCAI) toward the end of this pilot project. These individuals were selected by their societies to participate in this project and were provided with an overview of the methodology and the final priority scores generated by the CMTP workgroup. In general, the professional society representatives were supportive of the objectives and the results of the work. The priorities identified by the CMTP workgroup were generally consistent with the major evidence gaps that these professional association had identified when developing clinical guidelines for physicians.

Operationalizing the Research Priorities

During its final meeting, the workgroup discussed appropriate study designs for addressing the top five priorities (Table 2). Below we discuss some key points that emerged during the discussions for each of the questions. A wide variety of research methods, ranging from prospective trials to secondary analysis of existing data, were proposed to address the top priorities.

Research into the comparative effectiveness of PCI versus CABG for specific patient subgroups, ranked at the top of the final list. For female patients, specific subgroup analysis through pooling of individual level data from existing randomized clinical trials (RCTs) was recommended as the most appropriate research method, because, albeit female patients are recruited in trials, women-specific subgroup analyses are rarely undertaken and/or reported. For individuals over the age of 75 and ethnic minorities, individual observations are fewer and prospective randomized trials may be the most appropriate study type needed. However, for patients over 75, given possible difficulties in randomizing across treatment modalities in this subgroup, analysis of registry data (e.g., using the ACC and the STS databases) may be the only

feasible option. Because patients with CAD and left main disease or heart failure tend to be excluded from RCTs comparing different treatment modalities (the Surgical Treatment for Ischemic Heart Failure [STICH] trial [22] being an exception), new prospective studies may be the only way for addressing the existing uncertainty.

In addition to the comparative effectiveness of interventions for the treatment of CAD for specific subgroups, the role of noninvasive diagnostic testing and the development and use of adequate metrics to assess performance, were also ranked as top research priorities of relevance to decision makers.

To address the question of how the emergence of noninvasive tests is affecting outcomes, a stepwise approach, moving from observational studies to establish current practice and emerging trends to more experimental designs to assess the impact of different diagnostic pathways on management decisions and patient outcomes, would be required. The initial observational work could include a multicenter study to document utilization patterns across different centers, including tertiary and less specialized units supplemented by snapshots over time to assess whether utilization is going up. At the experimental stage, randomization across different management alternatives that would be defined in the initial observational stage would allow investigators to assess the impact of the newer diagnostics on outcomes.

Currently, in the field of CAD, there is a lack of reliable process measures that could serve as good surrogates for performance. Instead, volume of procedures and/or mortality rates are the main markers used at present, both of which are weak predictors of performance. To develop appropriate measures, first there needs to be a process of consensus building to identify and appropriately define surrogate markers (at the individual provider and institutional levels) based on evidence of recorded relationship to outcomes, supplemented by appropriate risk adjustment formulas. Overall, the group believed that, given the multiple factors affecting outcomes, it would be very challenging causally to attribute outcomes or performance to individual practitioners, even in the presence of appropriate risk adjustment. It was suggested that cost may be the best proxy for performance given these complexities.

DISCUSSION

Value of Prioritized List of Research Questions

As the list of new medical technologies continues to expand, so does the list of evidence gaps pertaining to their clinical effectiveness. To start addressing these questions, there is a need for a reproducible, systematic, and transparent method to prioritize the evidence gaps important to decision makers and translate them into research questions. This report describes one method of developing such a comparative effectiveness research agenda. A large part of the final meeting

of the group was devoted to discussing the strengths and weaknesses of this priority-setting process. Furthermore, we held one-to-one debriefing sessions with individual group members to further explore how the overall process could be improved. Below we discuss positive and negative aspects of the process as well as how this process could be strengthened and scaled up to inform a large, national CER initiative. The benefits of the process described above include:

Multistakeholder Involvement. A multistakeholder approach was adopted to ensure that a variety of perspectives are represented, ranging from frontline practitioners to researchers and methodologists, payers, device manufacturers, and patient representatives. Including individuals in their capacity as coming from different backgrounds helped make the discussions more productive and proves that different perspectives can converge in a setting of facilitated interaction.

Focus on a Single Disease Area/Condition. Since this pilot focused on one disease area, we were able to bring together experts, which enabled us to have more focused discussions throughout the process. Prioritizing evidence gaps is a time consuming process, of equal importance to evidence synthesis in a certain disease area and should be undertaken with a similar degree of time and resource commitment if it is to inform significant investment in research.

Up-to-Date Evidence. The evidence base is constantly evolving. Selecting a recently published systematic review from which to develop a research agenda makes it more likely that the final prioritized list of research questions reflect topical up-to-date evidence gaps of relevance to decision makers.

Transparency. Our structured approach to prioritizing research questions for CER means that the overall process and evidence considered at the different stages of the deliberations can be followed by external stakeholders, such as funders, policy makers, patients, and professionals. Any improvements or amendments in the process can also be assessed. Such transparency and accessibility are crucial given the important role CER is expected to play in improving current policy and practice in the United States.

User-Friendly Prioritization Method. Our approach was relatively inexpensive, reproducible, and applicable across different disease areas as long as the appropriate expertise is brought together in a facilitated environment.

The limitations of the process, as indicated by the workgroup participants include the following:

The Breadth of the Starting Point. A single comparative effectiveness review may not be broad enough to encompass the different types of question of importance for practice and policy. Even though we broadened the scope of the EPC report to include medical treatment, the issue of whether we included all the different treatment pathways in our analysis remains. Furthermore, during the pilot, the group identified two broad types of uncertainty: uncertainty about the effects of interventions and uncertainty about putting evidence that may already be available into practice. Our process dealt solely with the former type of uncertainty.

Presentation of Existing Information to Decision Makers. A significant challenge was making a clear accessible summary of the strengths and weaknesses of existing evidence on specific questions available to the workgroup. We developed formats that were designed to be accurate and convenient, but nonexperts in the workgroup still found it challenging to understand the key issues and gaps in the current evidence. One important refinement of the methodology described here would be to use a more systematic and structured approach (ideally tested with both experts and nonexperts) to presenting information to decision makers in the multistakeholder workgroup. This would ensure that all workgroup members understand current literature pertaining to the identified evidence gaps, as well as current clinical trials that may help to fill in those gaps.

The Validity of the Scoring Process. It was not clear whether all members of the workgroup were applying their scores in the same way. For example, the lay members' group tended to score everything very highly. Furthermore, a broader range of scores (e.g., 1–10 as opposed to 1–3) may have helped to better differentiate the different priorities. A larger scale could also allow for a larger distribution in the mean and median scores and would allow for a more accurate assessment of the level of group agreement.

The Medium of Interaction. Face to face meetings, possibly using a Delphi process, as opposed to relatively short 60-minute teleconference sessions may have been a more appropriate mode of interaction for the members of the workgroup and may also have clarified the validity concerns around scoring highlighted above.

Engagement of Lay Members. Given the technical nature of the project, it may have been better to provide additional support to lay members of the workgroup through individual sessions outside the workgroup meetings to clarify the objectives of every meeting and ensure members are comfortable with what is being asked of them.

Information Specialist Support. Given the importance of being aware of recently published or ongoing studies addressing the questions considered, it would have been useful to have an information specialist available to update the group members on relevant clinical trials or reports pertaining to the researchable questions being scored. This could take place real-time during the meetings and/or through preparing summaries of existing evidence to inform the prioritisation process, which could be provided as premeeting material. The latter was undertaken during this pilot with the help of the expert consultants and specialist workgroup members, albeit in a less systematic way.

Specificity of Questions. More focused research questions derived from the evidence gaps may have facilitated the prioritization process. The questions considered by the workgroup tended to be fairly compounded with multiple patient groups, interventions, and, often, outcomes. As noted above, we believe that a key area for refinement of our methodology would be in developing better techniques and formats to present the research questions to be scored by workgroup members.

Any future follow-up work would need to consider the issues outlined above. For example, including more technical support through an information specialist; working with a core group of experts earlier on in the process to help refine the key questions before putting them to the group for prioritization; and seeking broader public comment through interactive Web-based invited comment, may be points that could be relatively easily (and inexpensively) addressed and could potentially result in a significant improvement of the overall process. Finally, follow-up work will need to focus on how best to determine the optimal research methods to address each high priority research question and how to leverage a response in the form of a funding commitment on behalf of commercial and/or public sponsors of medical research.

In addition to refining the prioritization process, the greatest challenge lies in ensuring that this prioritized research agenda is disseminated and promoted, and that decision makers, researchers, and funding agencies are both empowered and willing to use the information generated through this process to inform policy decisions in the future.

Implications for the United States CER Effort

As CER is becoming increasingly important at both the federal and state levels in the United States, the issue of how a prioritized work program will be developed to guide investment in CER is finally getting the required attention from policy makers and academics alike. To date, no efficient, reliable, and scalable processes for identifying and prioritizing critical gaps in knowledge have been developed, although there have been several recent informative attempts by AHRQ, Medicare, and the IOM. Our project demonstrates that a multistakeholder process, underpinned by available evidence both in the form of RCTs or observational studies as well as in the form of expert views and deliberations among the stakeholders, can help generate a list of priorities for CER investment. However, the process we used was time-consuming and is in need of refinement before being applied more broadly across therapeutic areas. Disease-specific, multistakeholder panels like those used in this project could help narrow down the list of CER questions to those of direct relevance to decision makers, be it payers, patients, or professionals. An additional step of broad public Web-based electronic consultation would further ensure representativeness of the findings. Such panels could, for governance and impartiality

reasons, be convened by the IOM or by the Federal Coordinating Council for CER, as set out in the American Recovery and Reinvestment Act. However, their membership should be drawn from the public and private sectors and would ideally include broad stakeholder representation. For CER results to be useful, the most important first step is to ensure that the questions addressed reflect the evidence desired by all those who are actually making the clinical and health policy decisions.

Scaled up across major disease areas and broadened to involve as many expert stakeholders as possible, including patient and professional groups, this inclusive and transparent approach could inform national efforts to develop a work program for CER. Without this, the value of investing in CER will be hard to establish and buy-in from major stakeholders hard to gain.

SUPPLEMENTARY MATERIALS

Supplementary items 1–3 (www.journals.cambridge.org/thc)

CONTACT INFORMATION

Kalipso Chalkidou, MD, PhD (kalipso.chalkidou@nice.org.uk), National Institute for Health and Clinical Excellence, 71 High Holborn, London WC1V 6NA, UK
Danielle Whicher, MHS (Danielle.whicher@cmtpnet.org), Center for Medical Technology Policy, 400 E. Pratt Street, Suite 815, Baltimore, Maryland 21202
Weslie Kary, MPP, MPH (wkary@iha.org), Program Director, Medical Technology, Integrated Healthcare Association, 300 Lakeside Drive, Suite 1975, Oakland, California 94612
Sean Tunis, MD (Sean.Tunis@cmtpnet.org), Center for Medical Technology Policy, 400 E. Pratt Street, Suite 808, Baltimore, Maryland 21202

REFERENCES

1. Agency for Healthcare Research and Quality. *Effective health care*. http://effectivehealthcare.ahrq.gov/ (accessed July 3, 2008).
2. America's Health Insurance Plans Board of Directors. *AHIP board of directors statement on setting a higher bar for quality and value through comparative effectiveness research.* November 2007. http://www.ahip.org/content/default.aspx?bc=31|44|21474 (accessed July 3, 2008).
3. American Recovery and Reinvestment Act of 2009. 111th Congress, 1st Session. January 6, 2009. http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname = 111_cong_bills&docid=f:h1enr.pdf (accessed July 3, 2008). Washington, DC: US Senate.
4. Black N. A national strategy for research and development: Lessons from England. *Ann Rev Public Health*. 1997;18:484-505.
5. Bravata DM, Gienger AL, McDonald KM, et al. Systematic review: The comparative Effectiveness of percutaneous coronary interventions and coronary artery bypass graft surgery. *Ann Intern Med*. 2007;147:703–716.

6. Chilcott J, Brennan A, Booth A, et al. The role of modeling in prioritising and planning clinical trials. *Health Technol Assess*. 2003;7:235-251.

7. Committee on the NIH Research Priority-Setting Process. Institute of Medicine. *Scientific opportunities and public needs: Improving setting and public input at the National Institutes of Health*. Washington DC: National Academy Press; 1998.

8. Congressional Budget Office. *Research on the comparative effectiveness of medical treatments: Issues and options for an expanded Federal role*. December 2007. http://www.cbo.gov/ftpdocs/88xx/doc8891/12–18-ComparativeEffectiveness.pdf (accessed June 20, 2008).

9. Department of Health, United Kingdom. *Standing group on health technology: Report*. London: The Office of Information; 1994.

10. Donaldson MS, Sox HC. *Setting priorities for health technology assessment: A model process*. Washington DC: National Academy Press; 1992.

11. Institute of Medicine. *Compelling comparative effectiveness studies, priority assessment inventory project, February-September 2008*. http://www.iom.edu/CMS/28312/RT-EBM/55069/60576/60595.aspx (accessed June 20, 2008).

12. Institute of Medicine. *Knowing what works in healthcare: A roadmap for the nation*. Washington, DC: National Academies Press; 2008. www.iom.edu/CMS/3809/34261/50718.aspx (accessed June 20, 2008).

13. Kirschner N. *Improving availability of comparative effectiveness and cost-effectiveness information: An essential feature for a high-quality and efficient United States health care system*. A position paper of the American College of Physicians. Philadelphia: American College of Physicians; 2008.

14. Medicare Evidence Development and Coverage Advisory Committee. *Medicare evidentiary priorities*. August 2008. http://www.cms.hhs.gov/CoverageGenInfo/07_EvidentiaryPriorities.asp (accessed June 20, 2008).

15. Medicare Payment Advisory Commission. Producing comparative-effectiveness information. In: *Report to the Congress. Promoting greater efficiency in Medicare*. Washington, DC: Medicare Payment Advisory Commission; 2007. www.medpac.gov/chapters/Jun07_Ch02.pdf (accessed June 20, 2008).

16. Pharmaceutical Research and Manufacturers of America. *Profile 2008*. Washington, DC: PhRMA; 2008. http://www.phrma.org/files/2008%20Profile.pdf (accessed October 10, 2008).

17. Rowe G, Wright G, Bolger F. Delphi: A reevaluation of research and theory. *Technol Forecast Soc Change*. 1991;39:235-251.

18. Sackett DL, Hoey J. Why randomized controlled trials fail but needn't: a new series is launched. *CMAJ*. 2000;162:1301-1302.

19. Schoen C, Guterman S, Shih A, et al. *Bending the curve: Options for achieving savings and improving value in U.S. health spending*. The Commonwealth Fund, December 2007. http://www.commonwealthfund.org/usr_doc/Schoen_bendingthecurve_1080.pdf?section=4039 (accessed June 20, 2008).

20. United States Senate. Bill S.3408: *Comparative Effectiveness Research Act of 2008* (Introduced to Senate). July 31, 2008. Washington, DC: US Senate.

21. Van de Ven AH, Delbecq AL. Nominal versus interacting group processes for committee decision-making effectiveness. *Acad Manage J*. 1971;14:203-212.

22. Velasquez EJ, Lee KL, O'Conner CM, et al. The rationale and design of the Surgical Treatment for Ischemic Heart Failure (STICH) trial. *J Thorac Cardiovasc Surg*. 2007;134:1540-1547.

23. Vella K, Goldfrad C, Rowan K, Bion J, Black N. Use of consensus development to establish national research priorities in critical care. *BMJ*. 2000;320:976-980.